# Sur l'existence de la dérivée des fonctions monotones et sur quelques problèmes qui s'y rattachent.

Par Frédéric Riesz à Szeged.

## 1. Introduction.

Un des résultats les plus importants des recherches de M. Lebesgue sur la théorie de la mesure et l'intégration consiste en ce que toute fonction monotone admet presque partout une dérivée finie et déterminée. M. Lebesgue a établi ce théorème dans la première édition de son livre sur l'intégration, à la fin du dernier chapitre[1]), comme dernière conséquence de la théorie entière. Cependant ni l'idée de l'intégrale ni celle de la mesure n'interviennent dans l'énoncé du théorème ou tout au plus s'agit-il d'un cas particulier de cette dernière, l'expression „presque partout“ voulant dire que l'on n'aura à faire abstraction que d'un ensemble partiel de mesure nulle au plus, c'est-à-dire d'un ensemble qui peut être enfermé en un nombre fini ou une infinité dénombrable d'intervalles de sorte que la somme de ces intervalles soit arbitrairement petite. Malgré leur dénomination, les ensembles de mesure nulle s'introduisent et leurs propriétés essentielles[2]) s'établissent indépendamment de la théorie détaillée de la mesure.

---

[1]) H. Lebesgue, *Leçons sur l'intégration et la recherche des fonctions primitives*, Paris 1904, p. 128. En réalité, il s'y agit des fonctions à variation bornée, mais en vertu du théorème classique de Camille Jordan sur la décomposition de ces fonctions en fonctions monotones les deux énoncés sont équivalents. Il faut encore observer que l'auteur n'y parle que des fonctions continues; pour le cas des fonctions discontinues *voir* [4]) et la seconde édition de l'ouvrage cité, Paris 1928, pp. 186—188.

[2]) Il s'agira des propriétés suivantes. 1. *Le point est un ensemble de mesure nulle*; 2. *chaque ensemble faisant partie d'un ensemble de mesure nulle l'est également*; 3. *la réunion d'un nombre fini ou d'une infinité dénombrable*

Les premières démonstrations du théorème qui ne dépendent pas de la théorie de l'intégration, sont dues à M. FABER[3]) et à M. et M^me^ YOUNG.[4]) Ces démonstrations ainsi que celles inventées plus tard par divers auteurs sont basées, effectivement ou du moins en substance, sur des *théorèmes de choix* concernant des ensembles de points et des systèmes d'intervalles dont ils sont couverts; qu'il nous suffise d'en mentionner le théorème de M. VITALI.[5]) Dans ce qui suit, je vais développer une méthode différente, dans laquelle ces théorèmes de choix sont remplacés par un lemme appartenant aux éléments de l'Analyse.[6]) Outre son extrême simplicité, ce lemme a l'avantage qu'il permet d'enfermer les ensembles de mesure nulle dont il s'agira, en des systèmes de plus en plus petits d'intervalles par un procédé déterminé d'avance, pour ainsi dire entièrement mécanique.

Le lemme et la démonstration du théorème de M. LEBESGUE dont nous venons de parler, seront développés dans les paragraphes **2** et **3.** A leur suite, je profite de l'occasion pour exposer comment on déduit avec facilité du théorème démontré, en partie même immédiatement de notre lemme, plusieurs théorèmes bien connus et des plus importants appartenant à la théorie créée par M. LEBESGUE et concernant les courbes rectifiables, la dérivation des séries et la densité des ensembles. Je réserve pour une autre occasion d'exposer comment la même méthode s'applique à d'autres problèmes appartenant à cette théorie, en particulier à la dériva-

---

*de tels ensembles donne un ensemble de même espèce.* Les deux premières sont évidentes, la troisième vient comme il suit. Enfermons les ensembles en question en des systèmes d'intervalles de sorte que la somme de ces intervalles reste inférieure respectivement à $\varepsilon/2$, $\varepsilon/4$, $\varepsilon/8$, etc.; alors l'ensemble somme sera enfermé dans la totalité de ces intervalles, c'est-à-dire en un système d'intervalles dont la somme est plus petite que $\varepsilon$.

[3]) G. FABER, Über stetige Funktionen II., *Math. Annalen,* **69** (1910), pp. 372—443.

[4]) W. H. YOUNG and GRACE CHISHOLM YOUNG, On the Existence of a Differential Coefficient, *Proceedings London Math. Soc.,* (2) **9** (1910—11), pp. 325—335.

[5]) Pour les théorèmes de cette espèce, *cf.* T. H. HILDEBRANDT, The BOREL Theorem and its Generalizations, *Bulletin American Math. Soc.,* **32** (1926), pp. 423—474.

[6]) Je fus conduit à formuler ce lemme en étudiant un problème de MM. HARDY et LITTLEWOOD, *cf.* ma Note Sur un théorème de maximum de MM. HARDY et LITTLEWOOD, *Journal London Math. Soc.* (sous presse).

tion des fonctions d'intervalles, additives ou non, étudiée avec grand succès par MM. BURKILL et SAKS.

Observons encore, pour la commodité du lecteur, que les paragraphes **5, 6, 7** pourront être lus indépendamment l'un de l'autre.

## 2. Lemme fondamental.

*Étant donnée, dans l'intervalle* $a \leqq x \leqq b$, *une fonction continue* $g(x)$, *soit* $E$ *l'ensemble des points* $x$ *s'il y en a, intérieurs à l'intervalle et tels qu'il existe un* $x' > x$ *de sorte que* $g(x') > g(x)$. *Alors* $E$ *sera un ensemble ouvert c'est-à-dire qu'il sera formé d'un nombre fini ou d'une infinité dénombrable d'intervalles* $(a_k, b_k)$ *ouverts et disjoints, et pour chacun de ces intervalles, on aura*

$$(1) \qquad g(a_k) \leqq g(b_k).$$ [7]

En effet, comme les relations $a < x < b$, $x < x'$, $g(x) < g(x')$ ne sont pas troublées par les petites variations de $x$, l'ensemble $E$ sera ouvert. Envisageons un des intervalles ouverts et disjoints dont il est composé, soit $(a_k, b_k)$, et montrons que, pour tout $x_0$ intérieur à cet intervalle, on a $g(x_0) \leqq g(b_k)$; l'inégalité (1) s'ensuivra par continuité. A cet effet, désignons par $x_1$ le dernier point à droite tel que $x_1 \leqq b_k$, $g(x_1) \geqq g(x_0)$; nous aurons à montrer que $x_1 = b_k$. Or, si l'on avait $x_1 < b_k$, le point $x_1'$ qui correspond à $x_1$ conformément à l'hypothèse faite, devrait être situé au delà de $b_k$ et comme de plus, $b_k$ n'appartient pas à l'ensemble $E$, il viendrait que $g(x_1) < g(x_1') \leqq g(b_k) < g(x_1)$, ce qui implique contradiction.

Dans les applications qui suivent, nous aurons encore besoin d'une extension de notre lemme, cette extension se rapportant au cas des fonctions qui admettent des discontinuités de première espèce, c'est-à-dire telles que les valeurs limites $g(x-0)$ et $g(x+0)$ existent. Pour en simplifier l'énoncé, convenons de désigner par $G(x)$ la plus grande des valeurs $g(x-0)$, $g(x)$, $g(x+0)$ lorsque $a < x < b$ et par $G(b)$ la plus grande des valeurs $g(b-0)$, $g(b)$. Cela étant, *ceux des points* $x$ *intérieurs à* $(a, b)$ *qui donnent lieu à un* $x' > x$ *de sorte que* $g(x') > G(x)$, *forment un ensemble ouvert* $E$ *et l'on a*

$$g(a_k + 0) \leqq G(b_k),$$

[7] En réalité, c'est précisément le signe d'égalité qui a lieu, sauf peut-être si $a_k = a$.

$(a_k, b_k)$ *désignant l'un quelconque des intervalles ouverts et disjoints dont se compose l'ensemble* $E$.[8])

Cet énoncé général se démontre par un raisonnement analogue à celui qui précède; le lecteur se rendra compte des modifications évidentes qu'il aura à faire.

### 3. Existence de la dérivée des fonctions monotones.

Théorème de M. Lebesgue. *Toute fonction monotone dans un intervalle* $(a, b)$ *y admet presque partout une dérivée finie et déterminée.*

Pour fixer les idées, supposons que $f(x)$ ne soit jamais décroissante. Désignons de la manière usuelle par $\Lambda_d f(x)$ et $\lambda_d f(x)$ ou plus brièvement par $\Lambda_d$, $\lambda_d$ les nombres dérivés supérieur et inférieur à droite et par $\Lambda_g$ et $\lambda_g$ ceux à gauche. Pour démontrer le théorème, on n'aura qu'à prouver que

1. on a presque partout $\Lambda_d < \infty$;[9])
2. on a presque partout $\Lambda_d \leqq \lambda_g$.

En effet, en appliquant 2 à la fonction $-f(-x)$ il vient que

2'. on a presque partout $\Lambda_g \leqq \lambda_d$,

et en combinant 2 avec 2', on aura $\Lambda_d \leqq \lambda_g \leqq \Lambda_g \leqq \lambda_d \leqq \Lambda_d$; eu égard encore à la relation 1, cela nous assure que, presque partout, les quatre nombres dérivés admettent des valeurs finies égales entre elles, ce qu'il fallait démontrer.

Pour vérifier l'assertion 1, c'est-à-dire que l'ensemble $E_\infty$ des points $x$ pour lesquels $\Lambda_d = \infty$ est de mesure nulle, observons que cet ensemble est compris dans l'ensemble $E_C$ pour lequel $\Lambda_d > C$, $C$ désignant une quantité arbitrairement choisie. Or, la relation $\Lambda_d > C$ implique l'existence d'un $x' > x$ et tel que

$$\frac{f(x') - f(x)}{x' - x} > C,$$

c'est-à-dire que

$$f(x') - Cx' > f(x) - Cx$$

ou encore, en posant $g(x) = f(x) - Cx$, que $g(x') > g(x)$. Donc,

---

[8]) Le lemme est valable même pour toute fonction bornée $g(x)$; on n'aura qu'à définir la fonction $G(x)$ par ce qu'on appelle le maximum de la fonction $g(x)$ au point $x$ et à désigner par $g(a_k + 0)$ l'une quelconque des limites de $g(x)$ pour $a_k < x \to a_k$.

[9]) Au lieu de 1 il suffirait de prouver que $\lambda_d < \infty$.

en faisant abstraction, s'il est nécessaire, des extrémités $a$ et $b$ et des points de discontinuité de $f(x)$, c'est-à-dire d'un ensemble dénombrable donc de mesure nulle, l'ensemble $E_c$ fait partie de l'ensemble $E$ de notre lemme généralisé, se composant d'intervalles disjoints $(a_k, b_k)$ tels que $g(a_k+0) \leqq G(b_k)$. En observant encore que, dans notre cas, on a $G(x)=g(x+0)$ pour $x<b$ et $G(b)=g(b)$ et en introduisant de nouveau la fonction $f(x)$, il vient que

$$C(b_k-a_k) \leqq f(b_k+0)-f(a_k+0).$$ [10])

Il s'ensuit par addition que

$$C\sum(b_k-a_k) \leqq \sum[f(b_k+0)-f(a_k+0)] \leqq f(b)-f(a)$$

ce qui montre que, pour $C$ suffisamment grand, la longueur totale des intervalles $(a_k, b_k)$ sera aussi petite que l'on voudra. C'est-à-dire que l'ensemble $E_\infty$, compris dans ces intervalles excepté au plus un ensemble partiel de mesure nulle, devra être lui-même de mesure nulle. L'assertion 1 est donc démontrée.

Quant à l'assertion 2, soient $c<C$ deux quantités positives, et envisageons les points $x$ intérieurs à $(a, b)$ pour lesquels $\lambda_g < c$, en faisant abstraction, ici et dans la suite, des points de discontinuité de $f(x)$. En posant $g(x)=f(-x)+cx$ et en raisonnant comme tout à l'heure, on renfermera les points envisagés en des intervalles $(a_k, b_k)$, disjoints et tels que

$$f(b_k-0)-f(a_k-0) \leqq c(b_k-a_k).$$

Cela étant, on envisagera ceux des points de continuité, intérieurs aux intervalles $(a_k, b_k)$, pour lesquels $\Lambda_d > C$. En appliquant notre lemme cette fois à la fonction $g(x)=f(x)-Cx$ et cela séparément pour chaque intervalle $(a_k, b_k)$, on arrivera à enfermer les points en question dans des intervalles $(a_{kl}, b_{kl})$ disjoints, compris respectivement dans les intervalles $(a_k, b_k)$ et tels que

$$C(b_{kl}-a_{kl}) \leqq f(b_{kl}+0)-f(a_{kl}+0).$$

En répétant les deux procédés alternativement, on obtiendra une suite d'ensembles ouverts, chacun emboîté dans les précédents, et en désignant respectivement par $\Sigma_1, \Sigma_2, \ldots$ la longueur totale de

[10]) Pour $b_k=b$, on aura à remplacer $f(b+0)$ par $f(b)$ et de même on aura, dans ce qui suit, à mettre $f(a)$ au lieu de $f(a-0)$ et à faire aussi la convention analogue dans tous les cas quand on aura à appliquer le lemme à un intervalle partiel de $(a, b)$.

ces ensembles, il vient que

$$C\Sigma_2 \leqq \sum_{k,l} [f(b_{kl}+0)-f(a_{kl}+0)] \leqq \sum_k [f(b_k-0)-f(a_k-0)] \leqq c\Sigma_1,$$

c'est-à-dire que $\Sigma_2 \leqq {}^c/_C\, \Sigma_1$ et d'une façon générale que

$$\Sigma_{2n} \leqq {}^c/_C\, \Sigma_{2n-1}.$$

Il s'ensuit enfin que

$$\Sigma_{2n} \leqq ({}^c/_C)^n\, \Sigma_1 \to 0.$$

De cette façon, nous venons d'enfermer l'ensemble des points $x$ pour lesquels on a simultanément $\Lambda_d > C$ et $\lambda_g < c$, dans des systèmes d'intervalles dont la longueur totale tend en diminuant vers zéro. En réalité, nous avons négligé aussi, outre les points de discontinuité, les extrémités des intervalles dont sont composés nos systèmes; cependant, en tout, nous n'avons négligé qu'un ensemble dénombrable et par conséquent, les points $x$ pour lesquels $\Lambda_d > C$ et $\lambda_g < c$, forment un ensemble $E_{c,C}$ de mesure nulle.

Envisageons enfin les points $x$ pour lesquels $\Lambda_d > \lambda_g$, alors on pourra intercaler entre $\lambda_g$ et $\Lambda_d$ deux nombres positifs rationnels $c$ et $C$. C'est-à-dire que, en formant les ensembles $E_{c,C}$ pour toutes les valeurs rationnelles de $c$ et $C$, leur réunion $E^*$ contiendra tous les $x$ envisagés. Or, il n'y a qu'une infinité dénombrable de couples de nombres rationnels; par conséquent, l'ensemble $E^*$ est la réunion d'une infinité dénombrable d'ensembles de mesure nulle et, de cette sorte, il est de mesure nulle lui-même, ce qui achève la démonstration.

## 4. Théorème sur les suites de fonctions non décroissantes.

*Étant donné, sur l'intervalle $a \leqq x \leqq b$, une suite de fonctions $f_n(x)$ non décroissantes, supposons que les variations $f_n(b)-f_n(a)$ admettent zéro comme borne inférieure. Sous cette hypothèse, on a aussi, presque partout,*

$$\text{borne inf } f_n'(x) = 0. \tag{1}$$

Nous allons donner deux démonstrations de ce théorème, de nature entièrement différente. La première, basée sur notre lemme fondamental, montre que l'on a presque partout

$$\text{borne inf } \Lambda_d f_n(x) = 0$$

et cela sans tirer parti de l'existence de la dérivée elle-même. La seconde partira du théorème affirmant l'existence presque partout de la dérivée et elle a l'avantage d'être valable sous des hypothèses plus générales, comme par exemple pour les fonctions additives d'ensemble dans un espace à plusieurs dimensions et pour toute définition admise de la dérivée, bien entendu sous la condition que l'on ait démontré préalablement le théorème sur l'existence de la dérivée.

Voici la première démonstration. Soit $E^*$ l'ensemble des points $x$ pour lesquels la formule (1) est en défaut. L'ensemble $E^*$ est compris dans la réunion de l'infinité dénombrable des ensembles $E_m$ $(m=1, 2, \ldots)$, formés des points $x$ tels que $\Delta_d > {}^1/_m$ pour chaque $f_n(x)$. Donc il suffira de prouver que les $E_m$ sont de mesure nulle. Pour cela, soit $\varepsilon$ une quantité positive arbitrairement fixée et choisissons parmi les fonctions $f_n(x)$ une, soit $f_{n_0}(x)$, pour laquelle

$$f_{n_0}(b)-f_{n_0}(a)<{}^{\varepsilon}/_m;$$

grâce à l'hypothése faite, un tel choix sera toujours possible. Alors en appliquant à la fonction $f_{n_0}(x)$ le raisonnement employé au commencement du paragraphe précédent pour vérifier l'assertion 1, et en y posant $C={}^1/_m$, l'ensemble $E_m$ sera enfermé, sauf au plus une infinité dénombrable de points, dans des intervalles $(a_k, b_k)$ dont la longueur totale reste inférieure à $\varepsilon$. C'est-à-dire que l'ensemble $E_m$ est de mesure nulle.

La seconde démonstration se fera en extrayant de la suite des $f_n(x)$, une suite partielle $f^{(1)}(x), f^{(2)}(x), \ldots$ de sorte que l'on ait

$$f^{(n)}(b)-f^{(n)}(a)<{}^1/_{2^n}.$$

Cela étant, envisageons la série

$$s(x)=\sum[f^{(n)}(x)-f^{(n)}(a)];$$

elle est à termes non négatifs et converge uniformément. La fonction $s(x)$ sera monotone et l'on aura aussi évidemment

$$\sum_1^n \frac{f^{(n)}(x+h)-f^{(n)}(x)}{h} \leqq \frac{s(x+h)-s(x)}{h};$$

il s'ensuit que, pour tous les $x$ pour lesquels les dérivées de $f^{(i)}(x)$ et de $s(x)$ existent, on a

$$\sum_1^n (f^{(n)}(x))' \leqq s'(x).$$

Par conséquent, la série correspondante, composée des termes non négatifs $(f^{(n)}(x))'$, sera convergente presque partout, donc on aura presque partout $(f^{(n)}(x))' \to 0$ et, à plus forte raison,

$$\text{borne inf } f_n'(x) = 0, \qquad \text{c. q. f. d.}$$

## 5. Courbes rectifiables; formule de M. Tonelli.

Soient $x(t)$, $y(t)$, $z(t)$ des fonctions continues et à variation bornée dans l'intervalle $a \leq t \leq b$ et considérons la courbe rectifiable représentée par les équations $x = x(t)$, $y = y(t)$, $z = z(t)$. Soient de plus $a = t_0 < t_1 < t_2 < \ldots < t_n = b$, soient $P_0, P_1, P_2, \ldots, P_n$ les points de la courbe qui correspondent à ces valeurs du paramètre $t$ et désignons par $L$ la ligne polygonale $P_0P_1P_2\ldots P_n$ inscrite à la courbe $C$. Enfin, soit $P = P(t)$ un point variable de la courbe $C$ et soit $P' = P'(t)$ la projection orthogonale de $P$ sur la droite $P_{k-1}P_k$, correspondant à l'arc $P_{k-1}P_k$ sur lequel se trouve actuellement le point variable $P$. Désignons par $s = s(t)$ la longueur de l'arc $P_0P$ et par $l(t)$ celle de la partie $P_0P'$ de la ligne $L$, surtout en convenant, quant au calcul de $l(t)$, d'affecter la mesure du dernier segment $P_{k-1}P'$ de signe négatif au cas ou $P'$ se trouve à gauche de $P_{k-1}$ sur la droite respective $P_{k-1}P_k$. D'une façon plus précise, soit $l(t)$ la fonction continue s'annulant pour $t = 0$ et donnée à une constante additive près, par $\alpha x(t) + \beta y(t) + \gamma z(t)$ pour $t_{k-1} \leq t \leq t_k$ où $\alpha, \beta, \gamma$ désignent les cosinus directeurs du segment orienté $P_{k-1}P_k$, la valeur de la constante additive, variable avec $k$, étant définie par continuité. La fonction $l(t)$ ainsi définie restant encore indéterminée entre $t_{k-1}$ et $t_k$ lorsque $P_{k-1}$ coïncide avec $P_k$, nous convenons de poser dans ce cas $\alpha = 1$, $\beta = 0$, $\gamma = 0$.

Avec les conventions posées, la différence

$$d(t) = s(t) - l(t)$$

sera évidemment une fonction non décroissante de $t$ et sa variation $d(t_k) - d(t_{k-1})$ sur les intervalles $(t_{k-1}, t_k)$ sera égale à la différence de l'arc et de la corde respectifs en sorte que la variation $d(b) - d(a)$ sera donnée par la différence entre la courbe entière $C$ et la ligne $L$. La dérivée de la fonction $d(t)$ sera donnée, pour $t_{k-1} < t < t_k$, par

$$d'(t) = s'(t) - [\alpha x'(t) + \beta y'(t) + \gamma z'(t)], \tag{1}$$

$\alpha, \beta, \gamma$ désignant les cosinus directeurs respectifs et cela partout où les dérivées figurant au second membre existent. Or les fonctions à variation bornée $x(t), y(t), z(t)$ pouvant être décomposées, d'après le théorème classique de CAMILLE JORDAN, en fonctions monotones et la fonction $s(t)$ étant non décroissante elle-même, la relation (1) aura lieu presque partout dans l'intervalle $(a, b)$.

Observons enfin que, en appliquant l'inégalité de CAUCHY,

$$(2) \quad |\alpha x' + \beta y' + \gamma z'| \leqq [\alpha^2+\beta^2+\gamma^2]^{1/2}\,[x'^2+y'^2+z'^2]^{1/2} = [x'^2+y'^2+z'^2]^{1/2}$$

et que, d'autre part, on a évidemment

$$(3) \qquad x'^2+y'^2+z'^2 \leqq s'^2,$$

l'inégalité analogue concernant les accroissements finis des fonctions respectives n'exprimant autre chose que le fait que la longueur d'un arc n'est jamais dépassée par celle de la corde correspondante.

Des formules (1), (2), (3) on tire que

$$(4) \qquad 0 \leqq s'(t) - [(x'(t))^2+(y'(t))^2+(z'(t))^2]^{1/2} \leqq d'(t).$$

Cela étant, envisageons une suite indéfinie de lignes polygonales inscrites $L_m$ du type étudié et choisies de sorte que leur longueur tende vers celle de la courbe $C$. Alors on aura

$$d_m(b) - d_m(a) \to 0$$

pour les fonctions $d_m(t)$ qui y correspondent et par conséquent, en appliquant le théorème établi au paragraphe **4**, on a presque partout

$$\text{borne inf } d'_m(t) = 0.$$

Or, en comparant cette formule à l'inégalité (4) dans laquelle il est permis de remplacer $d(t)$, pour presque tous les $t$, par l'une quelconque des fonctions $d_m(t)$, il s'ensuit que l'on a précisément, presque partout,

$$(s'(t))^2 = (x'(t))^2 + (y'(t))^2 + (z'(t))^2.$$

Nous voilà arrivés à démontrer *la validité de la formule classique pour une courbe rectifiable du type le plus général même dans le cas d'un paramètre arbitraire.* Sous ces hypothèses générales, la formule est due à M. TONELLI.[11]) Elle se trouve dans la première

[11]) L. TONELLI, Sul differenziale dell'arco di curva, *Atti R. Accademia dei Lincei (Rendiconti)*, **25** (1º sem. 1916), pp. 207—213.

édition du livre de M. LEBESGUE avec des restrictions concernant le choix du paramètre $t$, qui pourtant n'affectent pas la généralité de la courbe. En particulier on y pourra lire la formule

$$(x'(s))^2 + (y'(s))^2 + (z'(s))^2 = 1,$$

cas particulier le plus important de la formule générale dont on déduit immédiatement l'existence de la tangente en presque tous les points de la courbe, c'est-à-dire presque partout par rapport à $s$, de même que la relation

$$\frac{\text{corde } MQ}{\text{arc } MPQ} \to 1,$$

valable pour presque chaque point $P$ da la courbe, l'arc $MPQ$ se contractant au point $P$.

Une autre conséquence immédiate de la formule de M. TONELLI est le fait suivant, d'ailleurs bien connu.

*La variation totale $T(t)$ de la fonction continue et à variation bornée $f(t)$, prise de $a$ jusqu'à $t$ et considérée comme fonction de $t$, a sa dérivée égale presque partout au module de $f'(t)$.*

En effet, il suffit de considérer la courbe $x = f(t)$, $y = 0$, $z = 0$, pour laquelle on aura évidemment $s(t) = T(t)$, ce qui fournit (presque partout) $(T'(t))^2 = (f'(t))^2$ et alors $T'(t) = |f'(t)|$.

La relation que nous venons de démontrer *est valable aussi pour les fonctions discontinues à variation bornée;* on la démontre dans ce cas par une légère modification du raisonnement qui précède. Cependant il est aussi facile de la déduire immédiatement du résultat obtenu au paragraphe précédent et cela, en appliquant ce dernier aux fonctions $f_n(t)$ correspondant à des décompositions $t_0 = a \leq t_1 \leq \ldots \leq t_m = b$ de plus en plus denses de l'intervalle $(a, b)$, égales, à des constantes additives près, à $T(t) \pm f(t)$ ou à $T(t)$ pour $t_{k-1} \leq t \leq t_k$, suivant le signe de la variation $f(t_k) - f(t_{k-1})$.

## 6. Séries de fonctions non décroissantes; théorème de M. Fubini.

Considérons la série

$$(1) \qquad \sum_1^\infty f_n(x),$$

formée de fonctions non décroissantes $f_n(x)$ définies pour $a \leq x \leq b$, et supposons qu'elle soit convergente pour $x = a$ et $x = b$. Alors la série (1) sera convergente dans tout l'intervalle $(a, b)$. En désignant sa somme par $S(x)$ et ses sommes partielles successives par $S_n(x)$, les restes $S(x)-S_n(x)$ de la série sont évidemment des fonctions non décroissantes et leur variation $S(b)-S_n(b)-S(a)+$ $+S_n(a)$ tend vers 0 avec $1/n$. Donc on aura presque partout, en vertu du théorème du paragraphe **4**,

$$\text{borne inf}\ (S'(x)-S'_n(x))=0,$$

ou, ce qui revient au même,

$$\text{borne sup} \sum_1^n f'_n(x) = \text{borne sup}\ S'_n(x) = S'(x).$$

En observant que $f'_n(x) \geqq 0$, il s'ensuit que l'on a presque partout

$$S'(x) = \sum_1^\infty f'_n(x).$$

On a donc le théorème de M. Fubini, d'après lequel *toute série convergente composée de fonctions non décroissantes peut être différentiée terme à terme.*[12])

## 7. Points de densité d'un ensemble; théorème de M. Lebesgue.

Soit $f(x)$ une fonction non décroissante, définie dans l'intervalle $a \leqq x \leqq b$ et soit $E$ un ensemble quelconque compris dans cet intervalle. Convenons de dire que la fonction $f(x)$ est *à variation nulle* sur l'ensemble $E$ lorsque la condition suivante est remplie. Pour chaque $\varepsilon$ positif et arbitrairement petit il existe un nombre fini ou une infinité dénombrable d'intervalles $(\alpha_k, \beta_k)$ de sorte que chaque point de $E$ soit intérieur à l'un au moins de ces intervalles et que, de plus, on ait

$$(1) \qquad \sum [f(\beta_k)-f(\alpha_k)] \leqq \varepsilon.$$

Je dis que *si la fonction $f(x)$ est à variation nulle sur l'ensemble $E$, on a $f'(x)=0$ presque partout sur l'ensemble $E$.*

---

[12]) G. Fubini, Sulla derivazione per serie, *Atti R. Accademia dei Lincei (Rendiconti),* **24** (1º sem. 1915), pp. 204—206. *Cf.* encore L. Tonelli, Successioni di curve e derivazione per serie, Nota II, *même Recueil,* **25** (1º sem. 1916), pp. 85—91; A. Rajchman, Sur la dérivabilité terme à terme des séries de fonctions monotones, *Fundamenta Math.*, **3** (1922), pp. 113—118 et p. 321; A. Rajchman et S. Saks, Sur la dérivabilité des fonctions monotones, *même Recueil,* **4** (1923), pp. 204—213.

En effet, posons $\varepsilon = 1/n$ et soit $\Sigma_n$ un système d'intervalles qui lui correspond au sens de la définition. Introduisons la fonction $f_n(x)$ comme il suit. Dans l'expression

$$\sum [f(\beta_k) - f(\alpha_k)],$$

supprimons les termes pour lesquels $x \leqq \alpha_k$ tandis que ceux pour lesquels $\alpha_k < x < \beta_k$ seront à remplacer par $f(x) - f(\alpha_k)$. La fonction $f_n(x)$ ainsi définie est évidemment non décroissante et l'on a manifestement

$$f_n(b) - f_n(a) \leqq 1/n.$$

En appliquant le théorème établi au paragraphe **4**, il s'ensuit que l'on a presque partout

(2) $$\text{borne inf } f'_n(x) = 0.\text{[13]}$$

Soit $x$ un point quelconque de l'ensemble $E$, soit $(\alpha_k, \beta_k)$ l'intervalle ou plus précisément l'un des intervalles appartenant au système $\Sigma_n$ et comprenant $x$ à son intérieur. Alors dans l'intervalle $(\alpha_k, \beta_k)$ la fonction $f_n(x)$ sera égale à

$$f(x) - f(\alpha_k) + g(x),$$

$g(x)$ se composant des mêmes termes que $f_n(x)$, excepté le terme $f(x) - f(\alpha_k)$ écrit à part. Donc $g(x)$ est une fonction non décroissante. Il s'ensuit que l'on a presque partout dans $(\alpha_k, \beta_k)$

$$f'_n(x) = f'(x) + g'(x) \geqq f'(x)$$

et par conséquent, on aura aussi

(3) $$f'_n(x) \geqq f'_r(x)$$

presque partout sur $E$ et cela pour chaque $n$. Une comparaison des formules (2) et (3) donne que

$$f'(x) = 0$$

presque partout sur $E$, ce qu'il fallait démontrer.

Soit maintenant $E$ un ensemble donné compris dans l'intervalle $(a, b)$ et désignons par $\mu(\alpha, \beta)$ la mesure extérieure de la partie de $E$ située entre $\alpha$ et $\beta$. On sait que la mesure extérieure d'un ensemble borné peut être définie en envisageant tous les systèmes formés d'un nombre fini ou d'une infinité dénombrable d'intervalles choisis de sorte que chaque point de l'ensemble soit

---

[13]) D'ailleurs il aurait suffi, pour le but envisagé, de se reporter aux formules analogues concernant $\Lambda_d f_n$ et $\Lambda_g f_n$, c'est-à-dire que, en réalité, on n'aura pas à supposer qu'on ait démontré préalablement l'existence des dérivées.

intérieur à l'un au moins des intervalles, puis en formant pour chaque système la somme des mesures des intervalles qui y appartiennent; cela étant, la mesure extérieure de l'ensemble sera définie comme la borne inférieure de toutes les sommes que l'on vient de former. Il découle immédiatement de cette définition que la mesure extérieure de l'ensemble somme d'un nombre fini ou d'une infinité d'ensembles ne peut jamais dépasser la somme des mesures extérieures de ces derniers et que de plus, $\mu(\alpha, \beta)$ est une fonction additive d'intervalle, c'est-à-dire qu'on a, pour $\alpha<\beta<\gamma$,

$$\mu(\alpha, \beta)+\mu(\beta, \gamma)=\mu(\alpha, \gamma).$$

Cela étant, envisageons la fonction

$$f(x)=x-\mu(a, x),$$

évidemment non décroissante. Je dis qu'elle est à variation nulle sur l'ensemble $E$. En effet, lorsqu'on enferme l'ensemble $E$ en des intervalles $(\alpha_k, \beta_k)$ de sorte que leur somme ne surpasse la mesure extérieure $\mu(a, b)$ de $E$ que de $\varepsilon$ au plus, on aura

$$\sum [f(\beta_k)-f(\alpha_k)]=\sum (\beta_k-\alpha_k)-\sum \mu(\alpha_k, \beta_k)\leq$$
$$\leq \mu(a, b)+\varepsilon-\sum \mu(\alpha_k, \beta_k)\leq \varepsilon.$$

Donc la fonction $f(x)$ est à variation nulle sur $E$ et par conséquent, en lui appliquant le théorème que nous venons d'établir, on aura

$$f'(x)=0$$

presque partout sur $E$. C'est-à-dire que

$$\text{(4)}\qquad \frac{\mu(x-h, x)}{h}\to 1, \quad \frac{\mu(x, x+h)}{h}\to 1, \qquad (0<h\to 0)$$

et cela pour presque tout point $x$ de l'ensemble $E$.

On appelle *points de densité* ceux des points de $E$ pour lesquels les relations (4) sont remplies. Avec cette dénomination et en observant que l'hypothèse que $E$ est borné, peut être levée immédiatement, on a le théorème: *Presque tous les points de l'ensemble E, de type général, en sont des points de densité.*

Ce théorème fut établi par M. Lebesgue dans le cas où $E$ est mesurable[14]; il lui a servi pour démontrer que toute fonction sommable est égale presque partout à la dérivée de son intégrale indéfinie, fait dont le théorème en question n'est qu'un cas très particulier. On sait d'ailleurs que si l'on admet la théorie de la

[14]) Cf. [1]), pp. 123—124.

mesure, l'énoncé général ci-dessus n'est qu'un simple corollaire du théorème de M. LEBESGUE; pour y passer, on n'aura qu'à observer que tout ensemble borné $E$ est compris dans un autre qui est mesurable et dont la mesure est égale à la mesure extérieure de $E$.[15]) L'intérêt principal que présente l'énoncé général consiste en ce que l'on peut s'en servir à maintes occasions à la place du théorème de M. LEBESGUE, sans avoir démontré préalablement que $E$ soit mesurable et même sans supposer connue la théorie de la mesure. Des démonstrations correspondant à ce dernier point de vue furent données par MM. BLUMBERG[16]) et SIERPIŃSKI.[17])

*(Reçu le 15 décembre 1931.)*

---

15) W. SIERPIŃSKI, Sur une extension de la notion de densité des ensembles, *Comptes rendus Acad. Sc. Paris,* **164** (1917), pp. 993—994.

16) H. BLUMBERG, A Theorem on Linear Point Sets, *Bulletin American Math. Soc.,* **25** (1918—19), pp. 350—353.

17) W. SIERPIŃSKI, Démonstration élémentaire du théorème sur la densité des ensembles, *Fundamenta Math.,* **4** (1923), pp. 167—171.